Optimal Ph D student–teacher ratio for Indian research universities: a production function approach

The notion of optimality pervades most domains of decision-making[1], including human development[2], resource allocation and optimality in human capital formation[3,4]. Some studies have estimated optimal school size with the view to enhancing student learning outcomes[5]. There have also been studies on the optimal amount of financial aid that is offered to students[6]. The notion of optimality should also pervade the higher education sector, as it places demands on scarce public funds and therefore needs rational rules for allocation.

India has one of the largest higher education systems in the world. It enrols the second highest number of students in the world after China[7]. Higher education is accessible only to a small segment of the Indian population. Only about 0.7% of those who join the Bachelor's programme enroll for a Ph D programme in India[8].

Over the years there has been a concern that India may be falling behind with regard to knowledge, the country's scientific output ranks with those of eastern Europe, like Albania[9].

In order to improve academic standards, the University Grants Commission (UGC) Regulations 2016 (ref. 10) has incorporated several provisions, including a restriction on the number of students that a teacher can guide at a time. Accordingly, a maximum of eight students (for Professors), six students (for Associate Professors) and four students (for Assistant Professors) per guide are permissible. This provision resulted in an animated debate in the academic circles and student protests on various university campuses. It is not known whether UGC had relied on any empirical assessment to arrive at these limits, but a limited search of the literature suggests a knowledge gap here. We address this issue and examine whether there is an optimal number of research students that an institution should have (per teacher).

A well-managed research programme would reflect on the research output of any higher education institution (HEI) and is measurable in terms of publications. Thus, our target variable would be total research output, but the total number of publications of a HEI size-dependent[11]. In order to overcome this problem and make the measure of research output scale neutral, we could normalize the size of an institution in terms of faculty size. Therefore, instead of using total output (i.e. total number of publications), we propose to use the average output per faculty (publication rate) as our target variable of interest.

The imposition of an upper limit suggests that the HEI regulator is implicitly operating in an optimality framework – what number of Ph D students per faculty produces the highest per capita research output.

Our framework is based on the well-established theory of optimization popular in economics. The production function approach demonstrates that under certain conditions, the firm maximizes output and any deviation from this would lead to sub-optimal results[1]. An educational institution in some senses is similar to a production firm. HEIs produce an output – knowledge, which can be evidenced from published research[12].

The production function approach, therefore, is suitable for our purposes at hand. It relates quantity (Q) produced to the various inputs of production (typically labour, L, and capital, K). Output per person (q) is dependent on capital (k) per unit of labour (eq. (1)).

$$q = \theta(k),\ q = \frac{Q}{L} = f\left(\frac{K}{L}, 1\right),$$
$$k = \frac{K}{L}, q = \theta\left(\frac{K}{L}\right). \quad (1)$$

We use the Cobb–Douglas version of the production function (eq. (2)) for its various desirable properties[1]

$$q = Ak^{\alpha}, \quad (2)$$

where A is the technology parameter and α is the proportion of output attributable to capital per unit of labour.

A log transformation (eq. (3)) allows us to linearize eq. (2).

$$\ln q = \ln A + \alpha \ln k. \quad (3)$$

The empirical analysis relies on the log–log econometric model as shown below

$$\ln q = \ln \hat{q} + \varepsilon = \ln A + \alpha \ln k + \varepsilon, \quad (4)$$

where $\ln \hat{q}$ is the predicted value of $\ln q$, and ε is the random error.

The term 'ε' in eq. (4) allows equalization of the predicted value $\ln \hat{q}$ to the actual value $\ln q$. The error term represents the other relevant variables, influencing $\ln q$, which are not explicitly included in our model[13].

We propose an empirical model based on the above theoretical framework. The target variable q (eqs (1) and (2)) is the average publication rate. We use a log transformation of the average publication rate (ln_Publication rate) as described above (eqs (3) and (4)) to estimate the 'β' coefficients[14]

$$\begin{aligned}(\text{ln\_Publication rate}) &= \beta_0 + \beta_1 (\text{ln\_Foundation year}) \\ &+ \beta_2 (\text{Total teachers})^2 \\ &+ \beta_3 (\text{Total teachers}) \\ &+ \beta_4 (\text{Annual Expenditure Per teacher}) \\ &+ \beta_5 (\text{Student\_PhD\_Teacher}) \\ &+ \beta_6 (\text{Student\_PhD\_Teacher})^2 \\ &+ \beta_7 (\text{Student\_PG\_Teacher}) \\ &+ \beta_8 (\text{Student\_PG\_Teacher})^2 \quad (5)\end{aligned}$$

The results of the model could be transformed suitably to predict the optimal number of Ph D students that a teacher should guide. If we differentiate the variable ln_Publication rate with respect to the variable Student_PhD_Teacher(s), we get

$$\frac{\partial[(\text{ln\_Publication rate})]}{\partial(\text{Student\_PhD\_Teacher})} = \beta_5 + 2\beta_6 (\text{Student\_PhD\_Teacher}). \quad (6)$$

The second-order differentiation yields

$$\frac{\partial^2[(\text{ln\_Publication rate})]}{\partial(\text{Student\_PhD\_Teacher})^2} = 2\beta_6. \quad (7)$$

We borrow the coefficient values from ref. 14 and find

$$\frac{\partial^2[(\text{ln\_Publication rate})]}{\partial(\text{Student\_PhD\_Teacher})^2} < 0, \quad (8)$$

suggesting the presence of a local 'maxima'.

Table 1. Optimal student–teacher ratio using coefficient values

Model	Target variable	Coefficient (β_5)	Coefficient (β_6)	Ratio $-(\beta_5/2\beta_6)$	Student_PhD_Teacher* (optimal)
1	ln_Publication rate	0.347	–0.0161	–(0.347)/(2*(–0.0161))	10.8
2	ln_Mod_x	0.569	–0.0293	–(0.569)/(2*(–0.0293))	9.7
3	ln_x	0.063	–0.036	–(0..063)/(2*(–0.036))	8.8

Source: Authors' calculation using the method of Mukhopadhyay *et al.*[14].

At the optimal (turning) point,

$$\frac{\partial[(\text{ln\_Publication rate})]}{\partial(\text{Student\_PhD\_Teacher})} = 0. \quad (9)$$

This implies that

$$\beta_5 + 2\beta_6\ (\text{Student\_PhD\_Teacher})^* = 0$$

and

$$(\text{Student\_PhD\_Teacher})^* = -\frac{\beta_5}{2\beta_6}. \quad (10)$$

Data on HEIs in India are available from multiple sources like the All India Survey of Higher Education (AISHE) and the National Assessment and Accreditation Council (NAAC). However, due to data limitations from these sources, the present analysis relies on data from the National Institutional Ranking Framework (NIRF). This database has two distinct advantages: (i) it is verified by NIRF, and (ii) it is in the public domain. We relied on NIRF data for 2017 and Scopus for publications.

We used three measures of publication rate according to Mukhopadhyay *et al.*[14] – one was a quantity measure, and two were exergy (quality) measures. The third measure Mod(x) in eq. (C) below standardizes the value of x in eq. (B) by size of total faculty.

Equation (A):

$$\text{Scopus per faculty} = \frac{\text{Total Scopus publications}}{\text{Total teachers}}.$$

Equation B:

$$x = \left\{\frac{\text{Publications in the top 25\% Scopus journals}}{\text{25(Total Scopus Publication)}}\right\}^2 \text{(Total Scopus Publication)}.$$

Equation C:

$$\text{Mod\_}x = \left\{\frac{\text{Publications in the top 25\% Scopus journals}}{\text{25(Total Scopus Publication)}}\right\}^2 \text{(Scopus Publication Rate)}.$$

While the performance index is represented by x (eq. (B)), the modified productivity index (termed Mod_x, is measured by dividing x above by the total number of teachers in each institution (eq. (C))[14].

The coefficient values from the regression results using the three models allow us to calculate the optimal PhD student–teacher ratio (Table 1).

Our results show that the research output (publication rate) could be maximized if there is an upper limit on the student–teacher ratio. Depending on whether we use the quantity measure or quality measure of publication rate (x and Mod(x)), the optimal PhD student–teacher ratio predicted by our model is in the range 8.8–10.8, *ceteris paribus*.

This has important implications for determining the number of students per teacher in the Indian HEIs. One limitation of this study is that it is based on a small, non-random sample of the 100 HEIs which feature in the top 100 list of NIRF 2017. The use of a larger database could further refine the coefficient values.

1. Varian, H. R., *Intermediate Microeconomics with Calculus: A Modern Approach*, W.W. Norton, New York, USA, 2014, p. 832.
2. Davies, A., *J. Socio-Econ*., 2009, **38**(2), 326–330.
3. Becker, G. S., *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, The University of Chicago Press, Chicago, USA, 1993, 3rd edn, p. 390.
4. Araujo, R. A., *Econ Lett.*, 2004, **85**(1), 71–76.
5. Fowler, W. J. and Walberg, H. J., *Educ. Eval. Policy Anal.*, 1991, **13**(2), 189–202.
6. Ehrenberg, R. G. and Sherman, D. R., *J. Hum. Resour*., 1984, **19**(2), 202.
7. Daugherty, L., Miller, T., Dossani, R. and Clifford, M., *Building the Links between Funding and Quality in Higher Education: India's Challenge*, Rand Corporation, Santa Monica, CA, USA, 2013, p. 52.
8. UGC, Annual Report 2016–17, University Grants Commission, New Delhi, 2017, p. 322.
9. Prathap, G., *Curr. Sci.*, 2017, **113**(3), 365–366.
10. UGC, University Grants Commission (minimum standards and procedure for award of M Phil/Ph D degrees) regulations, 2016. University Grants Commission, New Delhi, 2016; https://www.ugc.ac.in/pdfnews/4952604_UGC-(M.PHIL.-PH.D-DEGREES)-REGULATIONS,-2016.pdf
11. Mukhopadhyay, P., Sudarsan, P. K. and Tapaswi, M. P., *The Wire*, 2017; https://thewire.in/139651/nirf-framework-biased-favour-large-heis/
12. Griffiths, R., *Stud. High. Educ*., 2004, **29**(6), 709–726.
13. Wooldridge, J. M., *Introductory Econometrics: A Modern Approach*, South-Western Cengage Learning, Mason, OH, USA, 2013, 5th edn, p. 881.
14. Mukhopadhyay, P., Tapaswi, M. P., Sudarsan, P. K. and Sudarsan, K., *Curr. Sci.*, 2018, **114**(6), 1167–1173.

ACKNOWLEDGEMENT. We thank Gangan Prathap and an anonymous reviewer for helpful comments on an earlier version of this manuscript.

Received 13 September 2019; revised accepted 24 October 2019

PRANAB MUKHOPADHYAY[1]
P. K. SUDARSAN[1]
M. P. TAPASWI[2,*]

[1]*Department of Economics, and*
[2]*Visiting Faculty and Officer on Special Duty,*
Internal Quality Assurance Cell,
Goa University,
Goa 403 206, India
**For correspondence.*
e-mail: tapaswimurari@unigoa.ac.in